Neurocomputing 128 (2014) 466–475

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Sparse frontal face image synthesis from an arbitrary profile image

Lin Zhao [a], Xinbo Gao [a], Yuan Yuan [b,*], Dapeng Tao [c]

[a] School of Electronic Engineering, Xidian University, Xi'an 710071, China
[b] Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China
[c] School of Electronic and Information Engineering, South China University of Technology, GuangZhou 510640, China

ARTICLE INFO

Article history:
Received 17 January 2013
Received in revised form
29 May 2013
Accepted 20 August 2013
Communicated by M. Wang
Available online 22 November 2013

Keywords:
Face synthesis
Virtual frontal view
Inverse problems
Local patch

ABSTRACT

Frontal face image synthesis from an arbitrary profile image plays an important role in automatic video surveillance systems, and remains a challenge in computer vision. The strategies of partition are popular and promising for synthesizing frontal face images. However, conventional rectangular partition criterions fail to align corresponding patches in profile images and frontal face images. Given an arbitrary profile image, to synthesize a corresponding frontal face image which is smooth in texture and similar in appearance, we introduce a triangulation-based partition criterion and do synthesis based on sparse representation. The triangulation-based partition ensures the corresponding triangular patches are strictly aligned. And sparse representation adaptively finds the most similar patches for synthesis while abandons unlike patches. Furthermore, a confederate learning strategy is proposed to reduce the blocking artifacts caused by triangulation-based partition. Experimental results conducted on the Oriental Face database demonstrate the effectiveness of the proposed frontal face image synthesis method and advantages over previous works.

1. Introduction

Face synthesis remains an active topic of research in image processing and computer vision, due to its various potential applications in movie special effects manufacture, face recognition and video surveillance etc. Generally, face synthesis can be defined as synthesizing one person's face images given his/her face images in other poses. According to the pose in which the face image is given, the problem can be divided into two categories, i.e. synthesizing profile images from frontal images, and synthesizing frontal images from nonfrontal images. In the first case, because frontal faces contain more information and textures than profiles do, the synthesis is relatively easy. A common solution to do this is linear affine transformation, such methods like [1–3].

However, the synthesis of frontal images is left as an open problem. The reason why this problem is so difficult is that it is an inverse transformation of facial textures. The textures in a profile can be regarded as a deformation of the frontal face, and it can be directly described by affine transformations. But unfortunately, there is no transformation that is able to restore a frontal image from the deformed one. During the process of frontal face synthesis, we need to recover the textures occluded in the profile as well as repair the deformed textures. So it is a typical ill-posed problem to synthesize frontal images (see Fig. 1).

A synthesized frontal image should most likely have the following two properties: (1) **Nature:** the face image should be face-like; there should not be obvious deformity in local areas, and the facial textures need to be adequately smooth. (2) **Similarity:** the synthesized face must approximate the ground truth as much as possible; this means that it will be a failure if the synthesized image more closely resembles another person.

For this purpose, we present a method for frontal image synthesis, which fulfills the above two properties very well. For the first requirement, a triangulation based face division is utilized to make strict correspondence of facial textures, and then a confederate learning strategy is proposed to reduce the blocking artifacts among local textures. For the second requirement, sparse representation [4] is taken advantage of to model the similarities between local patches of face images. We test the proposed approach on the Oriental Face database [5]. Experimental results show that the proposed method provides better performance than previous works.

The rest of the paper is organized as follows. In Section 2, we review related work of frontal face image synthesis. In Section 3, an overview of the proposed algorithm is presented. We show a series of experiments investigating the performance of our approach in Section 4. Finally, the conclusion and future works are given in Section 5.

2. Related work

Many methods have been proposed to synthesize frontal images in the literature. In general, these methods can be divided into two categories. The first class builds a 3D face model from a

* Corresponding author. Tel.: +86 29 88889302.
E-mail addresses: sharling1988@gmail.com (L. Zhao), xbgao@mail.xidian.edu.cn (X. Gao), yuany@opt.ac.cn, yuan369@hotmail.com (Y. Yuan), dapeng.tao@gmail.com (D. Tao).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.08.018

single profile image and then the frontal image is obtained by using a weak perspective model. We term this one as geometric approach. A classic 3D model to do this is 3D Morphable Model (3DMM), which was proposed by Blanz et al. in [6]. 3DMM is trained via laser scanned 3D face datasets and allows linear variations of the shape and appearance. 3DMM based methods represent the state-of-the-art in face synthesis. The experimental results shown in [7] are not only excellent in image quality but also perfect in personal similarity. But the registration of a new individual to the model is quite slow. Because of its expensive computation, these methods are not currently suitable for real-time applications.

The other class is statistical approach, which is also the most common approach. These methods eschew domain-specific information about the 3D world, and hope to find the relationship between nonfrontal and frontal images via statistical learning. Without loss generality, the synthesis can be performed in two distinct ways. The first type of approaches synthesizes a frontal image as a whole. For example, Huang et al. [8] treated face synthesis as a manifold estimation problem, and the frontal image is synthesized by subspace reconstruction. Similarly, Prince et al. [9] proposed a probabilistic learning method to estimate the frontal image from a profile image. In the second type of approaches, a facial partition strategy is typically used. Faces are divided into many local patches, and each patch is synthesized separately. Chai et al. [10] presented a method named as linear local regression (LLR), in which faces were divided into many overlapped rectangular patches, and each patch is predicted by linear regression.

The statistical approaches introduced above have the advantages of relatively low computational cost and simplicity of implementation compared to the geometric approaches. However, the synthesized frontal images by these methods are not, in general, as good as for 3DMM based methods. For global statistical methods, it is difficult to obtain vivid local textures of facial organs such as eyes, nose and mouth. And for local statistical methods, the correspondence between local patches of frontal and non-frontal faces is hard to be guaranteed, which leads to poor results. In this paper, we develop a local statistical approach, which makes strict alignment of local patches and produces quite good results.

Before synthesizing frontal face images, head pose estimation and facial landmark localization are two commonly necessary steps. Fortunately, both of them have been widely studied in recent years and many effective methods have been proposed. Head pose estimation can be done by several kinds of approaches such as nonlinear regression methods [11,12] and manifold embedding methods [13,14]. A comprehensive survey of head pose estimation methods can be found in [15]. AAM (Active Appearance Model [16]) fitting [17,18] is a classic technique to do facial landmark localization. Nevertheless, these years more sophisticated algorithms [19,20] have been designed to locate facial landmarks on face images taken in the wild.

3. Frontal face image synthesis

3.1. Overview

The overall procedure of the proposed frontal face synthesis system consists of two main steps: fitting and synthesis (see Fig. 2). In the fitting step, the main purpose is to get the shape landmark points and extract the appearance image of the input profile (see Fig. 3). In the synthesis step, the extracted appearance image is divided into many different triangular patches via Delaunay triangulation [21]. Then, the textures in each corresponding triangle of the frontal face are synthesized via sparse representation [4]. Finally, a confederate learning strategy is performed to reduce the blocking artifacts.

In this paper, we assume that head pose estimation and facial landmark localization which are two techniques used in the fitting step are both solved using existing methods. For example, head

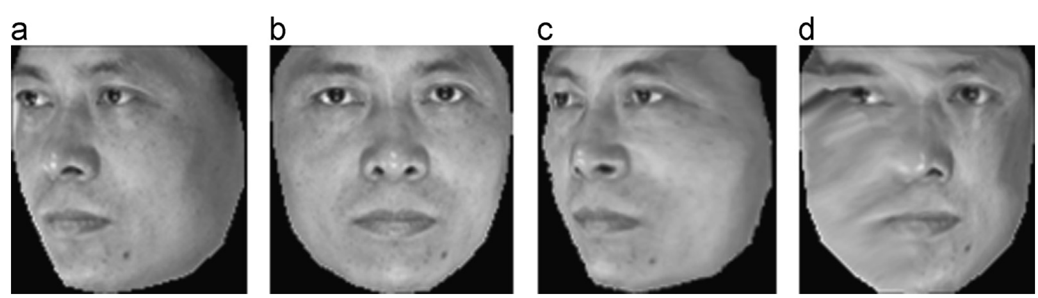

Fig. 1. Face synthesis results using linear affine transformation. Synthesized profile image (c) is quite excellent compared to the ground truth (a); however serious unnatural deformation occurs in the synthesized frontal image (d) compared to the ground truth (b).

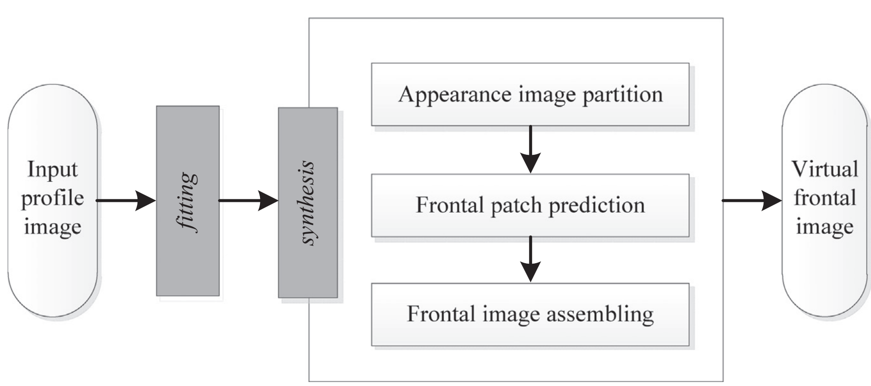

Fig. 2. Overall procedure of the proposed frontal face image synthesis system.